VETERANS HEALTH ADMINISTRATION
OFFICE OF PATIENT CARE SERVICES
TECHNOLOGY ASSESSMENT PROGRAM

BRIEF OVERVIEW:

SYSTEMATIC REVIEWS FOR SIMULATION TRAINING

Prepared by
Karen Flynn, DDS, MS
Program Manager

April 2009

TECHNOLOGY ASSESSMENT PROGRAM

An Effective Resource for Evidence-based Managers

VA's Technology Assessment Program (TAP) is a national program within the Office of Patient Care Services dedicated to advancing evidence-based decision making in VA. TAP responds to the information needs of senior VA policy makers by carrying out systematic reviews of the medical literature on health care technologies to determine "what works" in health care. "Technologies" may be devices, drugs, procedures, and organizational and supportive systems used in health care. TAP reports can be used to support better resource management.

TAP has two categories of products directed toward filling urgent information needs of its VA clients. TAP assigns a category to each new request based largely on the availability of studies from results of initial searches of peer-reviewed literature databases:

- The ***Short report*** is a self-contained, rapidly-produced qualitative systematic review of between 5 and 20 pages. It provides sufficient background information and clinical context to its subject technology to be accessible to a wide audience, including non-clinician managers.

- The ***Brief overview*** originated as an internal memo to VA clients with both well-defined and urgent information needs. It usually comprises 2 to 10 pages and assumes sufficient existing knowledge regarding clinical context and technology issues by its readers to omit these components of other TAP products. It often requires some additional reading of documents (provided with the overview for the client) to obtain a full and comprehensive picture of the state of knowledge on the topic.

All TAP products are reviewed internally by TAP's physician advisor and key experts in VA. Additional comments and information on this report can be sent to:

VA Technology Assessment Program • Office of Patient Care Services
Boston VA Healthcare System (11T) • 150 S. Huntington Ave. • Boston, MA 02130
Tel. (857) 364-4469 • Fax (857) 364-6587 • VATAP@va.gov

A SUMMARY FOR HTA REPORTS
Copyright INAHTA Secretariat 2001

VATAP is a member of the International Network of Agencies for Health Technology Assessment (INAHTA) [www.inahta.org]. INAHTA developed this checklist© as a quality assurance guide to foster consistency and transparency in the health technology assessment (HTA) process. VATAP will add this checklist© to its reports produced since 2002.

This summary form is intended as an aid for those who want to record the extent to which a HTA report meets the 17 questions presented in the checklist. It is NOT intended as a scorecard to rate the standard of HTA reports – reports may be valid and useful without meeting all of the criteria that have been listed.

Brief Overview:

Systematic Reviews Of Simulation Training

April 2009

Item	Yes	Partly	No
Preliminary			
1. Appropriate contact details for further information?	√		
2. Authors identified?	√		
3. Statement regarding conflict of interest?			√
4. Statement on whether report externally reviewed?		√	
5. Short summary in non-technical language?			√
Why?			
6. Reference to the question that is addressed and context of the assessment?	√		
7. Scope of the assessment specified?	√		
8. Description of the health technology?	√		
How?			
9. Details on sources of information?	√		
10. Information on selection of material for assessment?	√		
11. Information on basis for interpretation of selected data?	√		
What?			
12. Results of assessment clearly presented?	√		
13. Interpretation of the assessment results included?	√		
What Then?			
14. Findings of the assessment discussed?	√		
15. Medico-legal implications considered?			√
16. Conclusions from assessment clearly stated?	√		
17. Suggestions for further actions?	√		

ABBREVIATIONS IN THIS REVIEW

ASERNIP-S, Australian Safety and Efficacy Register of New Interventional Procedures-Surgical

BEME, best evidence for medical education

CME, continuing medical education

EDM, executive decision memo

EGD, esophagogastroduodenoscopy

LOS, length of stay

MAC, maximum arterial concentration

MAP, maximum arterial pressure

OPCS, Office of Patient Care Services

OSATS, Objective Structured Assessment of Technical Skills

OR, operating room

RCT, randomized controlled trial

RSI, rapid sequence induction

SMD, standardized mean difference

SEARC, simulation and innovation in education for health care (VHA national program under development)

SpO_2, percent oxygen saturation

VR, virtual reality

VT, video trainer

BRIEF OVERVIEW:

SYSTEMATIC REVIEWS FOR SIMULATION TRAINING

BACKGROUND

VHA's OPCS charged the VA Technology Assessment Program (TAP) with a review of the literature on simulation training as support for planning a National Program for Simulation Technology, Education and Research. TAP had already produced an overview ("review of reviews") of interventions for changing clinical behavior (Flynn 2007), which shares some citations with the current review. Together, the two overviews constitute a comprehensive picture of the best available research on the effectiveness of various strategies for clinical training and CME.

The EDM outlining the new Program defines simulation:

- *a) "Human simulation: simulated clinical environments with standardized patient actors;*
- *b) Mechanical patient simulators: artificial body parts and organs mimicking real-life patient scenarios.*
- *c) Virtual simulation; computer software and interactive web-based simulations.*
- *d) Haptic interfaces: robotic devices that mimic the feel and functionality of medical devices;*
- *e) Process modeling."*

The breadth and scope of TAP's implicit charge thus mandates another overview of available research based on existing systematic reviews. As explained in greater detail below, a catalog of systematic reviews provides an immediately accessible "snapshot" of the state of the research literature by highlighting those research questions for which a threshold quantity and quality of research, sufficient to warrant review effort, have been published. Such a catalog also synthesizes a larger body of literature than otherwise would be feasible for any single review, while defining gaps in the knowledge base for a research agenda.

"Computer-based training simulators have been used extensively, most notably in flight simulation. Over the past 20 years, surgical simulators have been developed, initially for training of minimally invasive surgery and more recently for open surgical simulation. The key effort in today's surgical simulation field is to develop metrics to evaluate how well the skills learned in a simulator translate to improvement in real surgical skills, execution of procedures, and team cooperation in the operating room. The American College of Surgeons has begun implementing a phased approach to introduce simulation in training and education for general surgery…" Rosen (2009).

"Acquisition of laparoscopic skills on simulators is becoming an essential part of surgical resident education. The Association of Program Directors and the American College of Surgeons in a collaborative effort recently published a national skills curriculum for use by training programs. Two of the 20 modules that comprise this national skills curriculum address the acquisition of laparoscopic skills…" Stefanidis (2009).

"Surgical training has traditionally been one of apprenticeship, where the surgical trainee learns to perform surgery under the supervision of a trained surgeon. Different procedures have different learning curves. Surgeons experienced in one procedure may not be experienced in another, and results for individual procedures improve with experience.

The different methods of laparoscopic surgical training include live animal training, human and animal cadaver training, training using a box trainer (also called a video trainer, VT), and virtual reality (VR) training (training using a computer simulation).

VT is currently being used widely for laparoscopic training, and has been shown to be better than standard training. VR has been reported to improve learning outcomes in different surgical procedures.

It also offers an ethical way of assessing the competency of a surgeon in performing a procedure, without risk to a patient.

Other reports suggest that VR training alone is inferior to traditional training for certain procedures. VR training has been used mostly for development of component skills (such as diathermy, clipping, suturing) and not training of an entire procedure. In contrast to the limited variability of data available during an aeroplane flight on which a pilot is trained using a custom designed simulator, anatomical variations are common in the human body, and skills acquired on a single computer simulation program may not be applicable in patients……

Although simulators can be expensive, traditional training is not without its costs. Junior surgeons take longer over operations than senior surgeons….average coats of this increased operating time were about $12,000US per year per resident during 1993-1997. The complication rate is also higher for junior surgeons. Thus the cost of the VR training system must be balanced against the cost of increased operating time and complication rates during traditional surgical training." Gurusamy (2008).

"Surgical and procedural skills training courses, delivered through dedicated skill laboratories and using various simulation-based approaches, are becoming accepted adjuncts to traditional patient-based training models. The fundamental assumption of simulation-based training is that the skills acquired in simulated settings are directly transferable to the operative setting. The effectiveness of individual simulators to teach procedural skills has yet to be proven…

Surgical training consists of developing cognitive, clinical, and technical skills, the latter being traditionally acquired through mentoring. Fewer mentoring opportunities have led to use of models, cadavers, and animals to replicate surgical situations and, more recently, to development of surgical skill centers or laboratories. However, the effectiveness of skills laboratories in teaching basic surgical skills (e.g., instrument handling, knot tying, and suturing) is not yet proven…" Sutherland (2006).

"Although simulation-based training has largely focused on technical skills, the acquisition of technical skills is only one aspect of surgical training. There is a range of nontechnical or human factors that are important for all-round surgical competence. It has been suggested that simulation allows for the development of the "pre-trained novice", an individual who has been trained to the point where many psychomotor skills and spatial judgments have been automated, allowing them to focus more on learning operative strategy and how to handle intra-operative complications, rather than wasting valuable operating room time on the initial refinement of psychomotor skills. With adequate pretraining, the trainee can gain maximum advantage from the supervised opportunities for training in the operating room or endoscopy suite…" Sturm (2008)

"The benefits of simulation derive from its standardization and reproducibility in contrast with the traditional apprenticeship approach to teaching where medical students and residents learn through practice with real patients in the clinic or hospital setting…shorter lengths of hospital stays, requirements for limited trainee work hours, and emphasis on patient safety, simulation has received greater attention…" McGaghie (2009).

METHODS

TAP first identified available systematic reviews of simulation or virtual reality training for health professionals of any discipline. TAP then updated searches conducted by review authors to confirm the presence or absence of subsequently published eligible studies that would change review conclusions.

Search strategy/selection criteria

TAP searched Medline and the Cochrane Library using the terms "simulation training" and "virtual reality", along with publication types (systematic review, meta-analysis) to identify systematic reviews published in English from 1990 to 2009 that synthesized research involving any health profession trainees preparing to treat adult human patients. Searches for subsequently published review-eligible RCTs were conducted on April 6, 2009 and all searches were finally updated on April 15, 2009.

Systematic reviews (detailed below) qualify as reproducible science and their production requires a threshold level of available primary research. Hence, a catalog of published systematic reviews provides an immediately accessible overview of the general status of a body of research literature. Conversely, the lack of published high-quality systematic reviews indicates a corresponding lack of published research on simulation issues of interest to the new VHA program.

TAP excluded:

- Narrative reviews;
- Reviews focused on technical development of simulation systems;
- Simulation training aimed at pediatric clinicians;
- Systematic reviews focused on uses of simulation other than training;
- Articles already included in systematic reviews;
- Primary studies where simulation was not the intervention or independent variable;
- "Quasi-systematic" reviews, i.e., those indexed or titled as systematic but which on close examination do not meet criteria or are inadequately reported to judge; these are noted but not abstracted in detail as their lack of rigorous methods risks significant bias.

One author (KF) selected citations for full-text retrieval, reviewed all articles, abstracted information, and prepared this overview.

ANALYTIC FRAMEWORK

Systematic reviews

Cook (1997) and Mulrow (1997) define systematic reviews: "*Systematic reviews are scientific investigations in themselves, with pre-planned methods and an assembly of original studies as their "subjects". They synthesize the results of multiple primary investigations by using strategies that limit bias and random error…"*

The same authors further specify characteristics of systematic reviews and contrast them with traditional narrative reviews: the latter synthesize articles without reporting methods of selection or quality assessment criteria and thus do not qualify as reproducible unbiased science.

Systematic reviews:

- Ask a focused clinical question;
- Conduct a comprehensive search for relevant studies using an explicit search strategy;
- Uniformly apply criteria for inclusion and exclusion of studies;
- Rigorously and critically appraise included studies;
- Provide detailed analyses of the strengths and limitations of included studies.

Systematic reviews can be quantitative (i.e., meta-analytic, applying statistical methods to summarize study results) or qualitative; in either case the inferences or conclusions of the review must follow logically from the evidence presented. The rigor of this approach is illustrated by the place of systematic reviews in evidence grading schemes (Cook 1995 and 1997; Guyatt 1995), where they receive the highest level designation. Reviews produced by the Cochrane Collaboration (www.cochrane.org) set the standard for rigor of methods and validity of conclusions. Cochrane reviews are meta-analytic where primary studies permit.

Some reviews classified by their authors or by indexing staff as "systematic" can be less than perfectly conducted and/or reported. Grimshaw (2002) critiques such reviews for:

- ignoring methodological weaknesses in primary studies, such as unit of analysis errors (analysis of unadjusted patient data when the unit of randomization is the physician), which results in artificially extreme p values and overly narrow confidence intervals;
- use of vote-counting methods, which add up the number of positive and negative comparisons and base effectiveness conclusions on the count. Positive comparison counts fail to provide an estimate of effect size and ignore the precision of the estimates from primary studies, or fail to exclude comparisons with unit of analysis errors.

While recognizing the limitations cited above, a vote count may be the logical response of an otherwise high-quality review to heterogeneity (in research questions, methods, interventions, or outcomes) among primary studies that precludes other methods of synthesis.

RESULTS

The six independent systematic reviews and two duplicate publications identified by TAP searches are outlined in Figure 1 below and abstracted in detail in Appendix Table 1. Four of the eight fully systematic reviews (Sturm, 2008; ASERNIPS, 2007; Sutherland 2006; Lynagh, 2007) represent stages of work by a group of Australian reviewers: their reference lists show substantial overlap. Although the AHRQ evidence review (Marinopoulos, 2007) is focused on CME, one of its key questions does specify simulation training.

As detailed in the figure and Appendix tables, high quality reviews cover over 300 studies of simulation training from database inceptions in the 1960s to the present. TAP identified very few subsequently published studies (Appendix Table 2) eligible for these reviews and none that would change review conclusions.

Quasi-systematic reviews are listed in the figure to acknowledge their authors' attempts to conduct reviews systematically and as an indication of a body of published literature, but are not formally included by TAP in the remainder of this overview.

Figure 1: Available systematic reviews for simulation or virtual reality training; covering published literature through 2008
Shading indicates related reviews: overlapping author lists or same review in different publication formats

Citation	Publication years covered/ number of studies included	Content
Systematic reviews		
Gurusamy (Cochrane; 2009a)	to March 2008: 23 studies	Virtual reality training for surgical trainees in laparoscopic surgery
Gurusamy (2008); print version of Cochrane review		
Schout (2008)	1980-2008: 45 studies	Simulation training models in urology
Sturm (2008): print version of ASERNIP-S (2007)	to 2006: 11 studies	• Transfer of simulation-based skills to the operating room • Effectiveness of surgical simulation compared to other methods of surgical training
ASERNIP-S (2007)		
Sutherland (2006)	To 2005: 30 studies	
Lynagh (2007)	1998-2006: 44 studies	Effectiveness of medical skills laboratories or simulators
Marinopoulos (2007)	2005-6 plus hand searching: 9 systematic reviews	Effectiveness of CME Key question 3: What is the evidence from systematic reviews about the effectiveness of simulation methods in medical education outside of CME?
Total **302 included studies with some duplication possible**		**6 fully systematic reviews; 2 duplicate publications.**

(Figure continued on next page.)

Citation	Publication years covered/ number of studies included	Content
Reviews with some characteristics of systematic methods but incomplete or uninterpretable reporting, insufficiently focused research questions, and/or inadequate quality assessment of included studies: quasi-systematic		
Tsang (2009)	Not reported	VR training for endovascular surgery
De la Rosette (2008)	Not reported	Methods for training and maintaining skills in percutaneous nephrolithotomy
Neequaye (2007)	Not reported	“the evidence for alternative tools currently available for endovascular skills training and assessment”.
Ravert (2002)	Not reported:	Computer-based simulation in the education process
	Total	**4 quasi-systematic reviews.**

SUMMARY/DISCUSSION

The substantial body of literature for simulation training is highly variable in quality and largely focused on minimally invasive procedures. As indicated by the quasi-systematic reviews listed in Figure 1 above, not even systematic reviews in this area of research are perfectly conducted or reported, and the most consistent comment of available fully systematic reviews is poor quality and heterogeneity of primary research.

The systematic review evidence suggests that virtual reality training for specific technical components of laparoscopic procedures is better than no training at all for novices prior to treating human patients. Less clear is the comparative role of simulation versus alternatives such as standard surgical apprenticeship training, other simulation approaches, or specific classroom methods like problem-based learning.

The literature thus raises more questions than it answers: professional associations including the American College of Surgeons have or are contemplating standardized curricula including simulation, but opinion plays a larger role than research evidence to date in precisely defining the role of this compelling technology in medical training.

Overriding concerns for the research agenda implicit in the list of ongoing trials below, should be more rigorous research, better reporting, and more careful editing by journals. The presence of randomized controlled trials and systematic reviews in this body of literature misleads the casual reader: the former often rely on convenience samples of uncertain relevance to target audience for training, are underpowered to detect important differences and generally lack other attributes of high quality research (specifically sample size calculations, reporting of randomization methods, and blinding of outcome assessors). The latter address unfocused questions, fail to tailor selection criteria to research questions, and can be too poorly reported to make optimal contributions to the knowledge base.

To borrow from two of the systematic reviews included here, and for which conclusions have not been changed by more recently published research:

“While there may be compelling reasons to reduce reliance on patients, cadavers, and animals for surgical training, none of the methods of simulated training (including computer simulation) has yet been shown to be better than other forms of surgical training. In addition, little is known about the real costs (including adverse outcomes in patients) of either simulated or standard surgical training.

“Adequately powered, well-designed and unconfounded RCTs (preferably multicenter with similar protocols) are needed and outcome assessors need to be blinded. Outcomes need to be tested in actual operative circumstances (or on validated systems). In particular, model simulation needs to be further tested against computer simulation. Studies of cost comparisons also need to be done. The RCTs dealt

exclusively with technical skills, although other skills such as cognitive skills and communication skills are clearly integral parts of surgical performance." Sutherland (2006)

"…The construct validity of the virtual reality (VR) simulators (ability to differentiate experienced versus inexperienced operators) has not to date been demonstrated for all simulated tasks. However, curriculum development on the basis of task validity and learning curve has been shown to improve performance during real procedures, and indeed, to shorten the time required to achieve proficiency in the real world…" Gurusamy (2009b).

IN-PROGRESS RESEARCH

Figure 2 (below) lists ongoing studies retrieved by searches at www.clinicaltrials.gov on April 1, 2009, using "simulation training". Figure 2 contents indicate the research agenda considered sufficiently compelling to warrant funding and organization of trials, along with issues likely to be resolved by trial results in the foreseeable future. Minimally invasive procedures continue as prominent subjects of current research although other areas have been added to the list.

Figure 2: In-progress studies

- **Retrieved from www.clinicaltrials.gov on April 1, 2009.**
- **Not listed: inactive, suspended, or withdrawn trials; trials in pediatric settings.**

Name/Purpose	Sponsor/location	Design/subjects	Estimated completion
Effectiveness of human simulation training for medical crisis management skills	University of Pittsburgh	RCT/ trainees at UPMC MCCTP	2007
Creating a proficiency-based virtual reality simulation training program for LAC	RCS/Ireland	RCT/ surgical trainees	2009
Trial of proficiency-based simulation training for general surgical trainees	RCS/Ireland	RCT/ junior surgical trainees	2009
Laparoscopic simulator training and its impact on surgical education	University of Texas Southwestern Medical Center	RCT/ gynecology residents	2008
High-fidelity simulation in health care education (critical care setting)	National Taiwan University Hospital	Before-and-after uncontrolled single group assignment (case series)	2011
Documenting a learning curve and test-retest reliability of a virtual reality training simulator in laparoscopic surgery	Columbia University	Case series/simulation-naïve undergraduate medical students	2005

LAC, laparoscopic assisted colectomy
MCCTP, multi-disciplinary critical care training program
RCS, Royal College of Surgeons
RCT, randomized controlled trial
UPMC, University of Pittsburgh Medical Center

APPENDIX

Table 1. Systematic reviews for simulation/virtual reality training (quasi-systematic reviews listed in Figure 1 not included)

Citation	Objective/Methods	Results, Conclusions, Recommendations, Comments
Gurusamy (2009a) Gurusamy (2008)	**Can virtual reality training supplement or replace conventional laparoscopic surgical training (apprenticeship)?** • Multiple databases and gray literature to March, 2008; • RCTs comparing virtual reality training Vs. other forms of training (video, no training, standard laparoscopic training, or different methods of VR training) in surgical trainees with little or no laparoscopic experience; • No language or publication status restrictions; • Data: characteristics of trials; methodologic quality; outcomes (morbidity, mortality, operating time, conversion rate, LOS);	**23 trials with 612 subjects:** • 4 trials (VR Vs video); 12 trials (VR Vs no training or standard training); 4 trials (VR, video, no training, standard training);3 trials (different methods of VR); • Most trials of high risk of bias; • In trainees without prior laparoscopic experience: VR decreased time to complete a task, increased accuracy, and decreased errors compared with no training; VR group was more accurate than video trained. • In trainees with limited laparoscopic experience: VR reduced operating time and error better than standard training. **Conclusions:** *"Virtual reality training can supplement standard laparoscopic surgical training of apprenticeship and is at least as effective as video trainer training in supplementing standard laparoscopic training. Further research of better methodological quality and more patient-relevant outcomes are needed"*
Sturm (2008) ASERNIPS (2007)	**Are skills acquired by simulation training transferable to the operative setting?** • Laparoscopic cholecystectomy and colonoscopy/sigmoidoscopy only; • Multiple databases to December, 2006; • RCTs reporting measures of task performance in simulation and in operative settings;	**10 RCTS; 1 nonrandomized comparative study:** • Mostly simulation in addition to standard training; • Overall performance: simulator-trained subjects scored higher in objective and subjective competence for initial assessment procedures; although in one study patient-trained subjects performed better; • Performance time: 4/5 studies found time improvement in simulator-trained subjects although one study found no difference; • Ability to complete procedure: simulator-trained subjects did better in 2/3 studies; one study found no difference; • Senior surgeon takeover: 6/8 trainees without simulator, 0/8 with; • Performance errors: 3 studies found significant reductions in errors with simulator training; • Time and motion: 2/3 studies reported improvements compared with no training; • Staff productivity: no significant difference in procedure volumes once simulator trained began patient colonoscopies; • Patient discomfort: ¾ studies reported less with simulator training; 1 study found no difference. **Conclusions:** *"Skills acquired by simulation-based training seem to be transferable to the operative setting. The studies included in this review were of variable quality and did not use comparable simulation-based methodologies, which limited the strength of the conclusions. More studies are required to strengthen the evidence base and to provide the evidence needed to determine the extent to which simulation should be a part of surgical training programs."*

Citation	Objective/Methods	Results, Conclusions, Recommendations, Comments
Schout (2008)	**Overview of training models for endo-urology and their validity):** • 1980-2008; • Descriptions of training models and/or their validity (face, content, construct, criterion validity defined).	**45 articles:** • 30 training models, 54 validation studies; • Largest number of models (9) in uretero-endoscopy; • 3 RCTs (7-136 subjects); • Criterion validity studies (impact on performance) scarce. **Conclusions:** *"Due to growing interest in training models in urology, it is increasingly urgent to determine which of these models are most valuable for postgraduate training. Because the validation studies published so far are few in number, have low evidence levels, and are composed of only a few RCTs, it is important that more randomized controlled validation studies including larger numbers of participants are performed."*
Lynagh (2007)	**Is performance in medical skills laboratories transferable to actual clinical performance and maintained over time?** • Multiple databases, 1998-2006; • Inclusion/exclusion: RCT evaluation of skills laboratory or simulator for medical education or skills training; undergraduate medical students or residents/interns; learner outcomes quantified with measure of procedural skill; full-text available in English.	**44 RCTs with 1602 subjects:** • 54% simulators for laparoscopic skills; 30% other procedures (endo-urological, urological, bronchoscopy, other general surgical skills); 16% evaluated other skill simulators (resuscitation, catheterization, trauma management, anesthesia, cardiac life support); • Approximately half of subjects were medical students, half post-graduates (residents, fellows, surgical trainees); 3 trials used mixed subject samples; • Range of skills laboratories (low-fidelity bench models high-fidelity/computer or VR) and intervention conditions (some studies randomizing to ≥ 3 groups); • Comparisons: 13 studies, simulation Vs no training; 12, simulation Vs standard training; 9, simulation Vs video box; 3 studies evaluated ≥ 2 simulators or their use under different conditions; 2 studies, model simulator Vs no training; 1, model Vs cadaver; 5, model Vs. standard; • Intensity and length of training: 10 minutes to 7 months; and many studies reported number of training sessions or number of times tasks repeated rather than length of practice time. **Effects of training:** • 70% of studies: simulator training significantly improved procedural skills Vs no or standard training; • 45% assessed transfer of performance to clinical skills (8 in animal models); and 2 studies assessed skills maintenance (both at 4 months). **Conclusions:** *"Medical skills laboratories do lead to improvements in procedural skills compared with no standard or no training at all when assessed by simulator performance and immediately post-training. However, there is a lack of well designed trials addressing the crucial issues of transferability to clinical practice and retention of skills over time. Further research must be carried out to address these matters if medical skills laboratories are to remain an integral component of medical education."*
Marinopoulos	**AHRQ Evidence review**	**9 reviews:**

Citation	Objective/Methods	Results, Conclusions, Recommendations, Comments
(2007)	**Key question 3:** **What is the evidence from systematic reviews about the effectiveness of simulation methods in medical education outside of CME?** • Multiple databases: 1990-2006.	• 8/9 evaluated simulation in skill acquisition (5, procedural or surgical skills; 2, communication skills; 2, physical exam skills); 2 reviews evaluated knowledge acquisition; • Reviews not completely reported to allow replication of results and covered wide variety of simulation methods. • **VR for surgical skills:** significant decrease in time to perform tasks and trend to decreased error rate; • **VT:** not superior to standard training or no training; insufficient evidence for computer Vs video; • **Training in GI endoscopy:** flexible sigmoidoscopy can be applied for clinical training of residents and fellow for patient comfort only; insufficient evidence to support simulators in GI endoscopy to improve patient outcomes. • **Teaching physical examination:** standardized patients to teach breast exam associated with better performance (improved ability to detect lumps or better post-test score); • **Teaching communication skills:** better empathy for patients with cancer and better communication of bad news; • **Knowledge acquisition:** effective in teaching physiologic principles to medical students (effect size pooled over 33 studies, 0.63 for use of computer simulation, but large range of effect sizes. **Summary:** *"Overall the direction of evidence points to the effectiveness of simulation training, especially in psychomotor skills (procedures or examination techniques) and communication skills, but the strength of evidence was considered low, due to the small number of appropriate studies, the scarcity of quantitative data, and other limitations. Several factors may be responsible for the inadequate quality of evidence in support of this method. In our view the most important factor is the lack of widely-accepted and standardized methods to quantify competency in procedural or communication skills. In addition, the high cost of simulation methods and difficulty in introducing clinical realism in a simulated environment are other factors that may be responsible for inadequate quality of evidence in this field."*
Sutherland (2006)	**What is the effectiveness of surgical simulation compared with other methods of surgical training?** • **Does repeated use improve performance?** (instructional effectiveness); • **Does the simulator measure the skills that it is designed to measure?** (construct validity); • **Is there a positive effect on patient outcomes?** (ultimate validation). • Multiple databases to April 2005; • RCTs assessing any training technique using at least some element of simulation Vs other methods of training or no training and reporting measures of surgical task performance;	**30 trials with 760 participants:** • Computer simulation generally better than no training, but not convincingly superior to standard training (surgical drills) or to video simulation (particularly when assessed by surgical performance); • Video simulation did not show consistently superior results to no training; • There were not enough data to determine if video simulation is better than standard training or the use of models; • Model simulation may be better than standard training, and cadavers may be better than models. **Conclusions:** *"While there may be compelling reasons to reduce reliance on patients, cadavers, and animals for surgical training, none of the methods of simulated training has yet been shown to be better than other forms of surgical training."*

Citation	Objective/Methods	Results, Conclusions, Recommendations, Comments
	• Participants: surgeons; surgical residents; medical students, others.	

Table 2: Subsequently published studies eligible for systematic reviews in Table 1

Citation	Objective/Methods	Results/Conclusions
Hallikainen (2009)	**RCT:** • Teaching anesthesia induction to undergraduate (year 4) medical students in Finland; • Full-size manikin Vs supervised teaching in OR; • Post-training assessment on simulator model with standardized evaluation list.	**46 students randomized:** • Passed test: 33% of traditional group, 87% simulation group; statistically significant differences in request for glycopyrrolate monitoring (P<0.001), SpO_2 monitoring (P<0.001), use of gloves in placing IV cannula (P=0.012), intubation attempt within 30 seconds (P<0.04), anesthesia gas set at MAC at least 1 (P<0.04), instructed anesthetic nurse to keep SpO_2 at least 95% (P<0.05), keep MAP at least 60 mmHg(P < 0.05), keep heart rate > 50/minute (P<0.002), keep end tidal PcO_2 4-5.5kPa (P<0.002). **Conclusions:** *"The simulation group performed better in 25% of tasks and similarly in the others compared with the traditional teaching group. With the same time and amount of teaching personnel we trained five or six students in the simulator compared with one in the operating theatre. Further research will reveal whether these promising results with simulation may be applied more generally in anaesthesiology teaching to medical students."*
Wenk (2009)	**RCT:** • simulation Vs problem-based learning (RSI in anesthesia); • 4th yr medical students attending anesthesia course; • Self-assessment questionnaire: knowledge and readiness for RSI at that point in the course; • Simulator testing10 days after training	**32 students:** Simulation group: significantly higher self assessment scores, but only slightly better theoretical and post-training simulation scores with moderate effect size (0.52). **Conclusions:** *"The current study demonstrates that both problem-based and simulator training lead to comparable short-term outcomes in theoretical knowledge and clinical skills. However, undesirably, simulator students overrated their anticipated clinical abilities and knowledge improvement."*
Kanumuri (2008)	**RCT:** • VR Vs computer-enhanced training for laparoscopic skills (suturing and knot tying) in novices; • 3rd year medical students; • Pre-training assessment on porcine model; • 4 weeks of training on VR (8 students) or computer-enhanced device (8); followed by post-training assessment.	**16 undergraduate medical students:** • Performance of groups comparable before and after training; • improvement over pretraining: task completion (94% Vs 18%; P<0.00); time (181 ±58 Vs 292 ± 24); • 88% of subjects thought haptic cues important in simulators; • Both groups agreed that devices were effective training tools; • Computer-enhanced group more likely to rate system as representative of reality (P<0.01): **Conclusions:** *"Training on virtual reality and computer-enhanced devices had equivalent effects on skills improvement in novices. Despite the perception that haptic feedback is important in laparoscopic simulation training, its absence in virtual reality did not impede acquisition of skill."*
Lucas (2008)	**RCT:** • VR laparoscopy simulator Vs no training; • Inexperienced medical students performed a baseline VR cholecystectomy, then were randomized (6 x 30 min	**32 undergraduate medical students, years 1 and 2:** • All completed study; • Groups comparable at baseline; • Post-training test: VR group performed significantly better on OSATS (27.9± 2.2 vs. 17.6 ±6.2;

Citation	Objective/Methods	Results/Conclusions
	unsupervised VR sessions Vs no training; • retested after training	P<0.001); • Trained students improved scores by at least 20% in each category (P<0.001); untrained improved only in "knowledge of procedure" (25%; P = 0.03) **Conclusions:** *"Skills training on a LAP Mentor VR simulator improved VR surgical performance. Before incorporating this simulator into resident education, the LAP Mentor will have to undergo testing for predictive and construct validity."*
Shirai 92008)	**RCT:** • Basic training (fundamentals of endoscopy) in esophago-gastroduodenoscopy: simulator +bedside Vs bedside training alone; • Hospital residents; • 5 hrs with GI-Mentor II; • Each subject then performed two endoscopies for assessment with 11-item 5-grade scale by 2 supervising physicians (blinding to training group not reported.	**20 residents randomized:** • 2 assessment procedures each x 11 items = n of 220; • No differences in demographics between groups; • No differences in mean score between two raters; • Evaluation scores significantly higher in simulator group, items for: insertion into esophagus; through pyloric ring; examination of duodenal bulb and fornix; • No significant difference in total procedure time; **Conclusions:** *"The performance of endoscopy was improved by 5 h of simulator training. The simulator was more effective with regard to the items related to manipulation skills. Computer-based simulator training in EGD is useful for beginners."*
Stefanidis (2008)	**RCT:** skill retention in OR following completion of proficiency-based laparoscopic skills simulator (Fundamentals of laparoscopic surgery suturing model**):** • N = 15 novices randomized to training or no training control • Assessment: simulator and live porcine laparoscopic Nissen fundoplication at training completion and 5 months	• **Training to proficiency:** 4.7 ±1.32 hrs and 41 ± 10 repetitions; • Trained subjects outperformed controls; • slight deterioration between post-test and retention test on simulator (505±22 Vs. 462±50, respectively; P<.05), • but not in OR (263±138 Vs. 279±88; P = .38). **Conclusions:** *"Proficiency-based simulator training results in durable in durable improvement in operative skills of trainees even in the absence of practice for up to five months. Minute simulator performance changes do not translate to the operating room."*

REFERENCES

ASERNIP-S. Surgical simulation for training: skills transfer to the operating room. Royal Australasian College of Surgeons/Australian Safety and Efficacy Register of New Interventional Procedures – Surgical. July 2007.

Cook, DJ, Guyatt GH, Laupaucis A, Sackett DL, Goldberg RJ. Clinical recommendations using levels of evidence for antithrombotic agents. *Chest.* 1995;108(4 Suppl):227S-230S.

Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions*. Annals of Internal Medicine.* 1997;126:376-380.

De la Rosette JJMCH, Laguna MP, Rassweiler JJ, Conort P. Training in percutaneous nephrolithotomy: a critical review. *European Urology.*2008;54:994-1003.

Flynn K. Systematic Reviews of Strategies for Changing Clinician Behavior to Improve Patient Outcomes. VHA Technology Assessment Program. April 2007.

Grimshaw JM, Eccles MP, Walker AE, Thomas RE. Changing physicians' behavior: What works and thoughts on getting more things to work. *Journal of Continuing Education in the Health Professions.* 2002;22:237-243.

Gurusamy KS, Aggarwal R, Palanivelu L, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database of Systematic Reviews.*2009a, Issue 1. Art NO.:CD006575. DOI:10.1002/14651858.CD006575.pub2.

Gurusamy K, Aggarwal R, Davidson BR. Authors' reply: systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. *British Journsl of Surgery.*2009b;96:221-3.

Gurusamy K, Aggarwal R, Panaivelu L, Davidson BR. Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. *British Journal of Surgery.*2008;95:1088-97.

Guyatt, GH, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. *Journal of the American Medical Association.* 1995;274:1800-1804.

Hallikainen J, Väsänen O, Randell T, Tarkkila P, Rosenberg PH, Niemi-Murola L. Teaching anaesthesia induction to medical students: comparison between full-scale simulation and supervised teaching in the operating theatre. *European Journal of Anaesthesiology.*2009;26:101-4.

Kanumuri P, Ganai S, Wohaibi EM, Bush RW, Grow DR, Seymour NE. Virtual reality and computer-enhanced training devices equally improve laparoscopic surgical skill in novices. *Journal of the Society of Laparo-endoscopic Surgeons.*2008;12;219-26.

Lucas SM, Zeltser IS, Behsalah K, Tuncel A, Jenkins A, Pearle MS, Cadeddu JA. Training on a virtual reality laparoscopic simulator improves performance of an unfamiliar live laparoscopic procedure. *Journal of Urology.*2008;180:2588-91.

Lynagh M, Burton R, Sanson-Fisher R. A systematic review of medical skills laboratory training: where to from here? *Medical Education.*2007;41:879-87.

Marinopoulos SS, Dorman T, Ratanawongsa N, Wilson LM, Bashar BH, Magaziner JL, Miller RG, Thomas PA, Propkopowicz GP, Qayyum R, Bass EB. Effectiveness of continuing medical education. Evidence report/Technology Assessment No. 149 (Prepared by the Johns Hopkins Evidence-based Practice Center, under contract No. 290-02-0018.) AHRQ Publication No. 07-E006. Rockville MD: Agency for Heathcare Research and Quality. January 2007.

McGaghie WC, Siddall VJ, Mazmanian PE, Myers J. Lessons for continuing medical education from simulation research in undergraduate and graduate medical education. *Chest.*2009;135:62S-68S.

Neequaye SK, Aggarwal R, Van Herzeele I, Darzi A, Cheshire NJ. Endovascular skills training and assessment. *Journal of Vascular Surgery.*2007;46:1055-64.

Ravert P. An integrative review of computer-based simulation in the education process. *Computers, Information, Nursing.*2002;20(5):203-8.

Rosen JM, Long SA, McGrath DM, Greer SE. Simulation in plastic surgery training and education: the path forward. *Plastic and Reconstructive Surgery.*2009;123:729-38.

Schout BMA, Hedrikx AJM, Scherpbier AJJA, Bemelmans BLH. Update on training models in Endourology: a qualitative systematic review of the literature between January 1980 and April 2008. *European Urology.*2008;54:1247-61.

Shirai Y, Yoshida T, Siraishi R, Nakamura H, Harada T, Nishikawa J, Sakaida I. Prospective randomized study on the use of a computer-based endoscopic simulator for training in esophagogastroduodenoscopy *Hepatology.*2008; 23;1046-50.

Stefanidis D, Walters KC, Mostafavi A. Heniford BT. What is the ideal interval between training sessions during proficiency-based laparoscopic simulator training? *American Journal of Surgery.*2009;197:126-9.

Stefanidis D, Acker C, Heniford BT. Proficiency-based laparoscopic simulator training leads to improved operating room skill that is resistant to decay. *Surgical innovation.* 2008;15(1):698-73.

Sturm LP, Windsor JA, Cosman PH, Cregan P, Hewett PJ, Maddern GJ. A systematic review of skills transfer after surgical simulation training. *Annals of Surger.*2008 Aug; 248(4):690-79.

Sutherland LM, Middleton PF, Anthony A, Hamdorf J, Cregan P, Scott D, Maddern GJ. Surgical simulation: a systematic review. *Annals of Surgery.*2006;243:291-300.

Tsang JS, Naughton PA, Leong S, Hill ADK, Leahy AL. Virtual reality simulation in endovascular surgical training. *The Surgeon.*2008:214-20.

Wenk M, Wauric R, Schotes D, Wenk M, Gerdes C, Van Aken HK, Pöpping DM. Simulation-based medical education is no better than problem-based discussions and induces misjudgment in zel-assessment. *Advances in Health Science Education.* 2009;14:159-71.